

The Search for Host Genetic Factors of HIV/AIDS Pathogenesis in the Post-Genome Era: Progress to Date and New Avenues for Discovery

Bradley E. Aouizerat · C. Leigh Pearce ·
Christine Miaskowski

Published online: 8 January 2011

© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract Though pursuit of host genetic factors that influence the pathogenesis of HIV began over two decades ago, progress has been slow. Initial genome-level searches for variations associated with HIV-related traits have yielded interesting candidates, but less in the way of novel pathways to be exploited for therapeutic targets. More recent genome-wide association studies (GWAS) that include different phenotypes, novel designs, and that have examined different population characteristics suggest novel targets and affirm the utility of additional searches. Recent findings from these GWAS are reviewed, new directions for research are identified, and the promise of systems biology to yield novel insights is discussed.

Keywords GWAS · Genome-wide association study · Disease progression · Viral load set point · SNP · Single nucleotide polymorphism

Introduction

Natural history cohort studies of HIV-1 were initiated in the United States and Europe during the peak of the AIDS epidemic (circa 1990). At the time, genome search methods were limited to family-based approaches, relegating the search for host genetic factors to candidate genes studies informed by a limited understanding of the pathophysiology of HIV [1•]. Nonetheless, progress was made, and the identification of the co-receptors for viral entry [2, 3] quickly led to the identification of allelic variants of the primary co-receptor, CCR5, that was depleted in HIV-1-infected individuals [4] and the carriers of which displayed delayed progression once infected with the virus. This key discovery demonstrated that host genetic factors could influence the course of HIV infection, provide insights into mechanisms, and suggest therapeutic targets. For an excellent review, the reader is directed to An and Winkler [1•].

The turn of the century brought new assay platforms and analytic methodologies that would allow for whole genome searches to be performed efficiently in unrelated individuals. To date, 11 genome-wide searches of varying design have been performed [5, 6, 7•, 8–10, 11•, 12–14]. While the concordance among the genome-wide association study (GWAS) findings is striking, these similarities are due in part to the phenotypes and subgroups of individuals selected for study.

GWAS of HIV Pathogenesis

The search for novel host genetic factors that influence HIV pathogenesis has focused on a restricted and largely overlapping set of phenotypes. These phenotypes have

B. E. Aouizerat (✉)
Department of Physiological Nursing and Institute for Human
Genetics, University of California San Francisco,
San Francisco, CA 94143-0610, USA
e-mail: bradley.aouizerat@nursing.ucsf.edu

C. L. Pearce
Department of Preventive Medicine,
Norris Comprehensive Cancer Center, Keck School of Medicine,
University of Southern California,
Los Angeles, CA 90089, USA
e-mail: cpearce@usc.edu

C. Miaskowski
Department of Physiological Nursing,
University of California San Francisco,
San Francisco, CA 94143-0610, USA

logically and primarily focused on three clinically relevant end points and measurements that are summarized below. It follows that the susceptibility loci identified thus far participate in innate and adaptive immunity. All of the studies reviewed below are of high quality in terms of study designs, genomic data collection, and data analysis. Moreover, a compelling mechanistic rationale exists for each gene identified.

HIV RNA viral load at set point refers to the acute phase of the initial HIV infection when viral replication attains a steady-state. Though challenging to define given the typical follow-up period in most cohort studies (i.e., bi-annual visits), and the exclusion of many HIV-infected individuals who did not meet the inclusion criteria, the concordance among findings from GWAS studies to date were striking (Table 1). Specifically, three loci have mapped to the major histocompatibility locus on chromosome 6 and have been verified in every cohort examined for viral load set point to date. Human leukocyte antigen (HLA) P5 (HCP5), HLA class B (HLA-B), and HLA-C harbor protective alleles that were associated with lower viral load at set point [6, 7, 13]. The high degree of correlation (termed linkage disequilibrium [LD]) between HCP5 and HLA-B made their independent associations with viral load at set point difficult to disentangle due to the small number of cases with rare recombination events between these loci. However, the association originally mapped to the HCP5 locus was subsequently dissected [6, 15•]. These analyses suggest that the HCP5 locus is associated with higher viral load at set point and the HLA-B locus (primarily HLA-B*57) is responsible for the protective effects detected by the single nucleotide polymorphism (SNP) located in HCP5 [15•].

Disease progression, defined as the time from seroconversion until the point at which immunosuppression occurs (i.e., a CD4⁺ T-cell count less than 350 cells/mm³, initiation of highly active antiretroviral therapy [HAART]), is a clinical end point of considerable interest. The extremes of the distribution in terms of disease progression, rapid progressors (RP), and long-term non-progressors (LTNP) have been the focus of several genome searches. Aside from the associations previously identified with the related phenotype, viral load at set point HCP5, HLA-B, HLA-C and variation of the zinc ribbon domain containing 1 (ZNRD1) gene are associated with disease progression [6, 7, 12].

Three studies that employed unique study designs are described below that resulted in the identification of additional novel disease loci for disease progression. The first study sought to refine the LTNP phenotype by excluding elite controllers. Elite controllers differ from LTNPs in that they suppress RNA viral load at levels that are below the limit of detection. Exclusion of elite controllers in a GWAS for LTNP uncovered an additional

risk locus, C-X-C chemokine receptor type 6 (CXCR6), validated in several cohorts [11•]. The second study sought to capture the entire spectrum of progression by initially screening three subgroups (i.e., RP, moderate progressors, LTNP), followed by replication in a larger cohort [8]. Variation in the prospero homeobox 1 (PROX1) gene was associated with slower progression. The third study first performed a two-stage linkage analysis in two family-based cohorts of macaques, and replicated an association signal detected on the X chromosome with viral load at set point and disease progression in a cohort of HIV-infected individuals [14]. The association signal mapped to an intragenic SNP located between the gene encoding for ribosomal protein S6 kinase alpha-6 (RPS6KA6) and the gene encoding for cyclin-1 (CYLC1). Subsequent validation in a larger sample may allow the gene underlying this association to be definitively identified.

Whereas the association signals with LTNP show considerable overlap with loci detected using disease progression [6, 7, 8, 12] as the phenotype, analysis of RP yielded unique loci [10] that have proven difficult to replicate in other cohorts. This inability to replicate may be due in part to the under-representation of RP in most cohorts. However, the possibility that individuals who are RP or LTNP harbor risk alleles that are unique to each tail in the distribution cannot be discounted. The minor alleles of SNP mapping to the gene-encoding protein arginine methyltransferase 6 (PMRT6), the gene encoding the sex-determining region Y-box 5 (SOX5) gene, and the gene encoding for the transforming growth factor, beta receptor associated protein 1 (TGFBAP1) alleles were depleted in RP. The risk allele mapping to the retinoid X receptor gamma (RXRG) gene was enriched in RP.

The majority of genome searches performed to date have focused on European-descent populations [5, 6, 7, 8, 10, 11, 12]. This approach is reasonable as it reflects the demographic of the epidemic when the cohorts analyzed thus far were initiated. The recent development of methods to account for more complex population substructure has paved the way to examination of populations with more diverse ancestry, such as Africans [9] and African Americans [13]. Although the first recently reported GWAS for viral load set point in African Americans failed to identify risk loci that exceeded the significance thresholds required of genome-wide searches, the associations with HCP5 and HLA-C were validated [13]. Examination of a completely different phenotype, maternal-to-child transmission in a cohort of HIV-serodiscordant children of HIV-infected mothers from Malawi, yielded several positional candidate genes. However, none exceeded the a priori significance thresholds. Further examination of these suggestive association signals may provide insights into the host genomic influence of the vertical transmission of HIV. Both studies

Table 1 Genome-wide searches for HIV-related traits

Trait	Sample description	Locus information	Associated effect	Study
• Viral load set-point	Caucasian seroconverters ^a	HCP5 (rs2395029)	Protective (lower viral load set point)	Fellay et al. [7•], 2007
• Disease progression	Initial GWAS: <i>n</i> =486	HLA-C (rs9264942)	Protective (lower viral load set point)	
	Focused replication: <i>n</i> =140	ZNRD1 (rs3869068)	Harmful (disease progression)	
• Plasma viral load (RNA)	Caucasian seroconverters ^b	HCP5 (rs2395029)	Protective (lower RNA & DNA viral load)	Dalmaso et al. [5], 2008
• Cellular viral load (DNA)	Initial GWAS: <i>n</i> =605	HLA-C (rs10484554)	Harmful (higher RNA & DNA viral load)	
	Focused replication: <i>n</i> =45	DDX40/YPEL2 (rs6503919) SDC2 (rs2575735)	Protective (lower RNA & DNA viral load)	
			Protective (lower RNA & DNA viral load)	
• Long-term non-progressors (LTNP)	Caucasian seroconverters ^{a, c}	HCP5 (rs2395029)	Protective (enriched in LTNP, lower viral load)	Limou et al. [12], 2009
	Initial GWAS: <i>n</i> =275	HLA-C (rs10484554)	Harmful (depleted in LTNP, higher viral load)	
	Focused replication: <i>n</i> =626	ZNRD1 (rs8321)	Protective (enriched in LTNP)	
• Rapid progressors (RP)	Caucasian seroconverters ^{a, c}	PRMT6 (rs4118325)	Protective (depleted in RP)	Le Clerc et al. [10], 2009
	Initial GWAS: 85 RP compared to 275 LTNP	SOX5 (rs1522232)	Protective (depleted in RP)	
		RXRG (rs10800098)	Harmful (enriched in RP)	
		TGFBAP1 (rs1020064)	Protective (depleted in RP)	
• Cellular susceptibility to HIV	Caucasian seroconverters ^d	LY6 (rs2572886)	Harmful (higher RNA viral load)	Loeuillet et al. [16], 2008
	Initial GWAS: <i>n</i> =254 cell lines			
	Focused replication: <i>n</i> =805			
• Viral load set-point	Caucasian seroconverters ^{d, e}	HCP5/HLA-B-5701 (rs2395029)	Protective (lower viral load set point, slower progression)	Fellay et al. [6], 2009
• Disease progression	Initial GWAS: <i>n</i> =1,397	HLA-C (rs9264942)	Protective (lower viral load set point, slower progression)	
	Focused replication: <i>n</i> =1,157			
• Disease progression	Caucasian seroconverters ^{a, h}	RPS6KA6/CYLC1 (rs5968255)	Protective (slower progression)	Siddiqui et al. [14], 2009
	Linkage scan: <i>n</i> =264 macaques			
	Focused replication: <i>n</i> =805			
• Viral load set-point	African American seroconverters ^{d, g}	[Failed to exceed significance threshold]	N/A	Pelak et al. [13], 2010
	Initial GWAS: <i>n</i> =515			
• LTNP, excluding elite controllers	Caucasian seroconverters ^c	CXCR6 (rs2234358)	Harmful (depleted allele in LTNP)	Limou et al. [11•], 2010
	Initial GWAS: <i>n</i> =605 compared to HIV- controls (<i>n</i> = 697)			
	Multiple replications: <i>n</i> =1,028			
• Mother-to-child transmission	African children	[Failed to exceed significance threshold]	N/A	Joubert et al. [9], 2010
	GWAS: 100 HIV+ compared to 136 HIV- children			
• Disease progression	Caucasian seroconverters ^{e, f}	PROX1 (rs17762192/rs1367951/rs17762150)	Protective (slower progression)	Herbeck et al. [8], 2010
	Initial GWAS: <i>n</i> =156			
	Focused replication: <i>n</i> =590			

^a Euro-CHAVI consortium^b ANRS PRIMO^c ARNS Cohort^d Swiss HIV cohort^e MACS^f Mixed^g DoD HIV NHS^h CHAVI consortium.

suggest that novel phenotypes may provide additional novel genes that influence other facets of HIV transmission and pathogenesis.

To date, two genome searches have pursued novel HIV traits. The first involved examination of not only circulating RNA viral load, but viral DNA that serves as an estimate of the HIV viral reservoir [5]. In addition to the verification of the previous associations with HCP5 and HLA-C, two additional associations with both lower RNA and DNA viral load were identified. The first was in the syndecan 2 (SYND2) and the second was with an intragenic SNP that detected two flanking positional candidate genes: DEAH (Asp-Glu-Ala-His) box polypeptide 40 (DDX40) and the human homolog of yippee-like 2 (YPEL2) [16]. Future validation efforts may be able to identify which of the two genes (DDX40 or YPEL2) underlies this later association signal.

Opportunities for Future Research

Although the discoveries made thus far using genome-wide searches are clear, many opportunities remain for additional discoveries. Perhaps the most pragmatic opportunity lies in the secondary analysis of the currently available GWAS data. Three analytic methods are likely to yield additional insights: meta-analysis, focused gene–gene interactions (i.e., a specific type of gene–gene interaction termed epistasis), and pathway analysis [17]. Locus-specific meta-analyses [12], the examination of specific gene–gene interactions [6], and pathway analysis performed by Fellay and colleagues [6] have been pursued in a subset of the GWAS described. The availability of several new and imminent GWAS datasets suggests that a more in-depth series of analyses are possible. These types of analyses may uncover additional genetic associations that could not surpass the statistical significance thresholds in the component studies due to limited power.

The GWAS reported to date have focused primarily on European-descent male populations that reflect the demographics of the AIDS epidemic at the time that natural history cohorts were built. Whereas considerable value exists in studying these cohorts, the demographics of HIV disease have shifted. Women of color are at highest risk for new infection. Cohorts that represent this shift in demographics (i.e., non-European descent, women), such as the Women's Interagency HIV Study (WIHS) [18], are currently available. The examination of gender-specific or gender-modified genetic associations was observed for two of the novel loci discovered by GWAS to date [12, 14]. In addition, the emergence of natural history cohorts of HIV in non-European-descent populations, such as the Centre for the AIDS Programme of Research in South Africa (CAPRISA) [19], will allow for the examination of the influence of

different HIV subtypes in the genetic associations identified to date and that may result in the identification of novel associations. An important caveat to GWAS in populations of non-European descent is that genetic marker coverage in the current commercially available arrays is variable. This limitation was in evidence in the study of Loeuillet and colleagues, where a key risk allele was not tagged in a commonly used commercial array for genome-wide variation measurement [16]. Fortunately, the goal of the 1000 Genomes Project (1KGP, www.1000genomes.org) is to dramatically expand the catalog of variation for the next generation of GWAS search tools, with the goal of identifying nearly all variants that exist at any appreciable frequency in human populations.

Recent advances in DNA resequencing have revolutionized the fields of genetics and genomics. Without doubt, whole genome sequencing will eventually supersede the current GWAS approach (i.e., measuring relatively common sequence variations). The barriers to the application of this genome search tool by research groups of even modest resources include the cost, error profiles, and limitations of the new sequencing platforms that differ from traditional sequencing technologies, and foremost are the bioinformatic challenges (for a review see [20, 21]). However, deep re-sequencing of candidate gene regions with high prior index of suspicion, such as those identified in multiple independent GWAS and supported by functional studies, is a method that is currently tenable and suffers more modestly from the barriers identified above. This method currently serves as a powerful adjunct to GWAS and is useful in the identification of rare variants and/or sequence anomalies not currently captured in commercially available genotyping arrays [22].

The greatest frontier for the discovery of host genetic factors that influence the pathogenesis of HIV lies ahead. To date, the genome searches have focused primarily on plasma RNA viral load and disease progression as estimated by peripheral blood CD4+ T-cell count. Though the concordance of the findings has affirmed the value of studying these traits and outcomes, the genes that these phenotypes have implicated play a primary role in either innate and/or adaptive immunity. The utility of examination of novel phenotypes, including *in vitro* characterization, is evident [16]. Recent advances in the understanding of the molecular mechanisms of HIV proviral latency [23] may inform the selection of future phenotypic analyses. The longitudinal data accrued in most natural history cohorts to date are limited by participant burden and cost, with bi-annual visits being the most common time-interval. However, the modeling of more complex patterns of change over time [24] (e.g., J-shape curves for viral load following acute infection, CD4+ T-cell count decline trajectories, latent variable analysis) may provide phenotypes that are superior to those

examined to date. The availability of banked serial biological specimens in many cohorts for which GWAS were reported suggests that cost effective studies can be performed by the addition of novel phenotypes that can be coupled with currently available genome-wide genetic marker data.

The emergence of HAART has naturally led to the examination of phenotypes captured prior to HAART due to the considerable complexity and ongoing evolution of these therapies in terms of drug targets. However, variable responses to HAART as well as differences in adverse event profiles remain a fundamental barrier to the success of these treatments [25]. Recent identification of gene polymorphisms that predict hypersensitivity reactions to different HAART drugs suggests that this line of inquiry is tenable [25]. The recent development of resource-efficient drug exposure measures, such as the measurement of HAART deposition in hair, holds promise to provide additional insights into pharmacogenomic risk factors as confounding due to self-report adherence is largely circumvented [26–28].

GWAS of HIV-Associated Comorbid Disease

With the advent of HAART, HIV infection transitioned from an acute to a chronic disease. It is now well accepted that chronic HIV infection and/or HAART may influence several co-morbid diseases of aging as well as elicit disease-specific conditions [29, 30]. These conditions include, but are not limited to, neuropathy [31], nephropathy, atherosclerosis [32], metabolic perturbations [33], and neurocognitive disorders [34•]. In addition, the role of the host genome in the setting of co-infections that are common in HIV-infected individuals (e.g., hepatitis C, human papilloma virus) is an active area of research, particularly given the increased risk for common co-morbid disease (e.g., chronic kidney disease [35]).

The complex and poorly understood natural history of HIV infection necessitated rigorous longitudinal follow-up and in-depth characterization of the participants (e.g., demographic characteristics, clinical characteristics, co-morbid diseases, cell repository) enrolled into natural history cohorts. A natural byproduct of these intense and sustained studies is the possibility to not only contribute to an understanding of HIV and subsequently response to HAART, but also common co-morbid disorders and diseases. Below, two GWAS that serve as compelling exemplars of the importance of studying comorbid diseases in the setting of HIV are described.

Risk for chronic kidney disease is strikingly elevated in the setting of HIV infection. In addition, differences by race are observed with African Americans at increased risk. Mapping by admixture linkage disequilibrium, a method

that detects ancestral risk alleles for disease in groups of individuals from recently mixed populations, resulted in the identification of a novel risk locus (i.e., myosin, heavy chain 9, non-muscle [MYH9]) [36]. Although recent evidence suggests that an adjacent locus may underlie the association with the MYH9 locus [37], the power of the approach is clear.

An equally promising discovery by GWAS of atherosclerosis in the setting of HIV was recently reported by Shrestha and colleagues [38]. Their search resulted in the identification of two SNPs in tight LD associated with carotid intima-media thickness mapping to the ryanodine receptor (RYR3). Previous work not only implicated a role for these SNPs in the etiology of cardiovascular disease, but the RYR3 protein is also known to interact with the HIV Tat protein. Clearly, additional research is warranted to better understand the role of the host genome in risk for common disease within the context of HIV.

Systems Biology: A New Vista to Understanding HIV Pathogenesis

Clearly, the complex host–viral interactions that underlie HIV pathogenesis will not be unraveled by GWAS alone. The integration of several other components of both cellular and organism-level processes will be required, termed a systems biology approach. Recent examinations of different functional RNA (fRNA) classes and gene expression profiles of specific host immune cell populations have yielded unique insights into HIV infection [1••]. The examination and integration of genomic, epigenomic, transcriptomic, proteomic, metabolomic, and viral protein interactome are requisite. Though bioinformatic, computational, and statistical barriers to the integration of these data exist, new solutions to these challenges emerge daily. The integration of these data is sure to suggest novel therapeutic opportunities to interfere with the host–viral interaction to stymie effective infection.

Conclusions

The success of GWAS in the identification of HIV infection and control of viral levels (i.e., set point) is clear. However, the genes discovered to date and their variations only explain a portion of the variance in these traits. The addition of novel phenotypes in cohorts with pre-existing genome-wide data, the examination of novel cohorts by GWAS, and the application of novel analytic approaches and data mining will undoubtedly yield novel insights into HIV pathogenesis and risk for common co-morbid diseases. GWAS remains a cost-effective strategy to identify genes of

interest that can be the focus of more resource-intensive functional and molecular studies. And finally, a systems biology approach will permit the integration of GWAS findings with other facets of cellular and organism-level biology and to prioritize targets for future therapies.

Acknowledgments This work was funded by grants from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH) and NIH Roadmap for Medical Research (KL2 RR024130).

Disclosure No potential conflicts of interest relevant to this article were reported.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Papers of particular interest, published recently, have been highlighted as:
 - Of importance
 - Of major importance
- An P, Winkler CA. Host genes associated with HIV/AIDS: advances in gene discovery. *Trends Genet.* 26(3):119–131. *This is an excellent review of host genetic studies of HIV/AIDS. In addition to review of the first five GWAS of HIV/AIDS, the review provides an excellent overview of candidate gene studies and identifies areas for future research.*
- Deng H, Liu R, Ellmeier W, Choe S, Unutmaz D, Burkhardt M, Di Marzio P, Marmon S, Sutton RE, Hill CM, Davis CB, Peiper SC, Schall TJ, Littman DR, Landau NR. Identification of a major co-receptor for primary isolates of HIV-1. *Nature.* 1996;381(6584):661–666.
- Dragic T, Litwin V, Allaway GP, Martin SR, Huang Y, Nagashima KA, Cayanan C, Maddon PJ, Koup RA, Moore JP, Paxton WA. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature.* 1996;381(6584):667–673.
- Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM, Saragosti S, Lapoumeroulie C, Cogniaux J, Forceille C, Muyldermans G, Verhofstede C, Burtonboy G, Georges M, Imai T, Rana S, Yi Y, Smyth RJ, Collman RG, Doms RW, Vassart G, Parmentier M. Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature.* 1996;382(6593):722–725.
- Dalmaso C, Carpentier W, Meyer L, Rouzioux C, Goujard C, Chaix ML, Lambotte O, Avettand-Fenoel V, Le Clerc S, de Senneville LD, Deveau C, Boufassa F, Debre P, Delfraissy JF, Broet P, Theodorou I, Association AGW. Distinct genetic loci control plasma HIV-RNA and cellular HIV-DNA levels in HIV-1 infection: the ANRS Genome Wide Association 01 study. *PLoS One.* 2008;3(12):e3907.
- Fellay J, Ge D, Shianna KV, Colombo S, Ledergerber B, Cirulli ET, Urban TJ, Zhang K, Gumbs CE, Smith JP, Castagna A, Cozzi-Lepri A, De Luca A, Easterbrook P, Gunthard HF, Mallal S, Mussini C, Dalmau J, Martinez-Picado J, Miro JM, Obel N, Wolinsky SM, Martinson JJ, Detels R, Margolick JB, Jacobson LP, Descombes P, Antonarakis SE, Beckmann JS, O'Brien SJ, Letvin NL, McMichael AJ, Haynes BF, Carrington M, Feng S, Telenti A, Goldstein DB, Immunology NcHIV. Common genetic variation and the control of HIV-1 in humans. *PLoS Genet.* 2009;5(12):e1000791.
- Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, Zhang K, Gumbs C, Castagna A, Cossarizza A, Cozzi-Lepri A, De Luca A, Easterbrook P, Francioli P, Mallal S, Martinez-Picado J, Miro JM, Obel N, Smith JP, Wyniger J, Descombes P, Antonarakis SE, Letvin NL, McMichael AJ, Haynes BF, Telenti A, Goldstein DB. A whole-genome association study of major determinants for host control of HIV-1. *Science.* 2007;317(5840):944–947. *This is the first GWAS of HIV/AIDS. The study demonstrated the feasibility and utility of host genome searches.*
- Herbeck JT, Gottlieb GS, Winkler CA, Nelson GW, An P, Maust BS, Wong KG, Troyer JL, Goedert JJ, Kessing BD, Detels R, Wolinsky SM, Martinson J, Buchbinder S, Kirk GD, Jacobson LP, Margolick JB, Kaslow RA, O'Brien SJ, Mullins JI. Multistage genomewide association study identifies a locus at 1q41 associated with rate of HIV-1 disease progression to clinical AIDS. *J Infect Dis.* 201(4):618–626.
- Joubert BR, Lange EM, Franceschini N, Mwapasa V, North KE, Meshnick SR, the NcHIVAVI. A whole genome association study of mother-to-child transmission of HIV in Malawi. *Genome Med.* 2(3):17.
- Le Clerc S, Limou S, Coulonges C, Carpentier W, Dina C, Taing L, Delaneau O, Labib T, Sladek R, Group AG, Deveau C, Guillemin H, Ratsimandresy R, Montes M, Spadoni JL, Therwath A, Schachter F, Matsuda F, Gut I, Lelievre JD, Levy Y, Froguel P, Delfraissy JF, Hercberg S, Zagury JF. Genomewide association study of a rapid progression cohort identifies new susceptibility alleles for AIDS (ANRS Genomewide Association Study 03). *J Infect Dis.* 2009;200(8):1194–1201.
- Limou S, Coulonges C, Herbeck JT, van Manen D, An P, Le Clerc S, Delaneau O, Diop G, Taing L, Montes M, van't Wout AB, Gottlieb GS, Therwath A, Rouzioux C, Delfraissy JF, Lelievre JD, Levy Y, Hercberg S, Dina C, Phair J, Nonfield S, Goedert JJ, Buchbinder S, Estaquier J, Schachter F, Gut I, Froguel P, Mullins JI, Schuitemaker H, Winkler C, Zagury JF. Multiple-cohort genetic association study reveals CXCR6 as a new chemokine receptor involved in long-term nonprogression to AIDS. *J Infect Dis.* 202(6):908–915. *This is a GWAS of the LTNP phenotype that identified a novel but mechanistically sound locus for the trait. The refinement of the non-progression phenotype by exclusion of elite controllers highlighted the value of exploring other HIV/AIDS phenotypes and in applying new knowledge to refine phenotypes.*
- Limou S, Le Clerc S, Coulonges C, Carpentier W, Dina C, Delaneau O, Labib T, Taing L, Sladek R, Deveau C, Ratsimandresy R, Montes M, Spadoni JL, Lelievre JD, Levy Y, Therwath A, Schachter F, Matsuda F, Gut I, Froguel P, Delfraissy JF, Hercberg S, Zagury JF, Group AG. Genomewide association study of an AIDS-nonprogression cohort emphasizes the role played by HLA genes (ANRS Genomewide Association Study 02). *J Infect Dis.* 2009;199(3):419–426.
- Pelak K, Goldstein DB, Walley NM, Fellay J, Ge D, Shianna KV, Gumbs C, Gao X, Maia JM, Cronin KD, Hussain SK, Carrington M, Michael NL, Weintrob AC, Infectious Disease Clinical Research Program HIVWG, National Institute of A, Infectious Diseases Center for HIVAVI. Host determinants of HIV-1 control in African Americans. *J Infect Dis.* 201(8):1141–1149.
- Siddiqui RA, Saueremann U, Altmuller J, Fritzer E, Nothnagel M, Dalibor N, Fellay J, Kaup FJ, Stahl-Hennig C, Nurnberg P, Krawczak M, Platzer M. X chromosomal variation is associated with slow progression to AIDS in HIV-1-infected women. *Am J Hum Genet.* 2009;85(2):228–239.

15. •• Catano G, Kulkarni H, He W, Marconi VC, Agan BK, Landrum M, Anderson S, Delmar J, Telles V, Song L, Castiblanco J, Clark RA, Dolan MJ, Ahuja SK. HIV-1 disease-influencing effects associated with ZNRD1, HCP5 and HLA-C alleles are attributable mainly to either HLA-A10 or HLA-B*57 alleles. *PLoS One*. 2008;3(11):e3636. *Careful modeling resulted in the dissection of these two tightly linked loci and provided evidence of unique contributions of each locus. Of note, the HCP5 locus was associated with progression and not delayed progression, contrary to previous report due to the strong correlation with the protective HLA-B locus.*
16. Loeuillet C, Deutsch S, Ciuffi A, Robyr D, Taffé P, Munoz M, Beckmann JS, Antonarakis SE, Telenti A. In vitro whole-genome analysis identifies a susceptibility locus for HIV-1. *PLoS Biol*. 2008;6(2):e32.
17. Cantor RM, Lange K, Sinsheimer JS. Prioritizing GWAS results: A review of statistical methods and recommendations for their application. *Am J Hum Genet*. 86(1):6–22.
18. Barkan SE, Melnick SL, Preston-Martin S, Weber K, Kalish LA, Miotto P, Young M, Greenblatt R, Sacks H, Feldman J. The Women's Interagency HIV Study. WIHS Collaborative Study Group. *Epidemiology*. 1998;9(2):117–125.
19. van Loggerenberg F, Mlisana K, Williamson C, Auld SC, Morris L, Gray CM, Abdool Karim Q, Grobler A, Barnabas N, Iriogbe I, Abdool Karim SS, Team CAIS. Establishing a cohort at high risk of HIV infection in South Africa: challenges and experiences of the CAPRISA 002 acute infection study. *PLoS One*. 2008;3(4):e1954.
20. Kircher M, Kelso J. High-throughput DNA sequencing—concepts and limitations. *Bioessays*. 32(6):524–536.
21. Moore JH, Asselbergs FW, Williams SM. Bioinformatics challenges for genome-wide association studies. *Bioinformatics*. 26(4):445–455.
22. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. *PLoS Biol*. 8(1):e1000294.
23. Bisgrove D, Lewinski M, Bushman F, Verdin E. Molecular mechanisms of HIV-1 proviral latency. *Expert Rev Anti Infect Ther*. 2005;3(5):805–814.
24. Muthén B, Muthén LK. Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. *Alcohol Clin Exp Res*. 2000;24(6):882–891.
25. Vidal F, Gutierrez F, Gutierrez M, Olona M, Sanchez V, Mateo G, Peraire J, Vilades C, Veloso S, Lopez-Dupla M, Domingo P. Pharmacogenetics of adverse effects due to antiretroviral drugs. *AIDS Rev*. 12(1):15–30.
26. Bernard L, Vuagnat A, Peytavin G, Hallouin MC, Bouhour D, Nguyen TH, Vilde JL, Bricaire F, Raguin G, de Truchis P, Ghez D, Duong M, Perronne C. Relationship between levels of indinavir in hair and virologic response to highly active antiretroviral therapy. *Ann Intern Med*. 2002;137(8):656–659.
27. Duval X, Peytavin G, Breton G, Ecobichon JL, Descamps D, Thabut G, Leport C. Hair versus plasma concentrations as indicator of indinavir exposure in HIV-1-infected patients treated with indinavir/ritonavir combination. *AIDS*. 2007;21(1):106–108.
28. Gandhi M, Ameli N, Bacchetti P, Gange SJ, Anastos K, Levine A, Hyman CL, Cohen M, Young M, Huang Y, Greenblatt RM, Women's Interagency HIVS. Protease inhibitor levels in hair strongly predict virologic response to treatment. *AIDS*. 2009;23(4):471–478.
29. Mothe B, Perez I, Domingo P, Podzamczar D, Ribera E, Curran A, Vilades C, Vidal F, Dalmau D, Pedrol E, Negro E, Molto J, Paredes R, Perez-Alvarez N, Gatell JM, Clotet B. HIV-1 infection in subjects older than 70: a multicenter cross-sectional assessment in Catalonia, Spain. *Curr HIV Res*. 2009;7(6):597–600.
30. Volberding PA, Deeks SG. Antiretroviral therapy and management of HIV infection. *Lancet*. 376(9734):49–62.
31. Robinson-Papp J, Simpson DM. Neuromuscular diseases associated with HIV-1 infection. *Muscle Nerve*. 2009;40(6):1043–1053.
32. Farrugia PM, Lucariello R, Coppola JT. Human immunodeficiency virus and atherosclerosis. *Cardiol Rev*. 2009;17(5):211–215.
33. Stankov MV, Behrens GM. Contribution of Inflammation to Fat Redistribution and Metabolic Disturbances in HIV-1 Infected Patients. *Curr Pharm Des*.
34. • Levine AJ, Singer EJ, Shapshak P. The Role of Host Genetics in the Susceptibility for HIV-associated Neurocognitive Disorders. *AIDS Behav*. 2009;13(1):118–132. *This is an excellent review of HIV-associated neurocognitive disorders. An examination of plausible candidate genes is provided.*
35. Estrella MM, Fine DM. Screening for chronic kidney disease in HIV-infected patients. *Adv Chronic Kidney Dis*. 17(1):26–35.
36. Kopp JB, Smith MW, Nelson GW, Johnson RC, Freedman BI, Bowden DW, Oleksyk T, McKenzie LM, Kajiyama H, Ahuja TS, Berns JS, Briggs W, Cho ME, Dart RA, Kimmel PL, Korbet SM, Michel DM, Mokrzycki MH, Schelling JR, Simon E, Trachtman H, Vlahov D, Winkler CA. MYH9 is a major-effect risk gene for focal segmental glomerulosclerosis. *Nat Genet*. 2008;40(10):1175–1184.
37. Tzur S, Rosset S, Shemer R, Yudkovsky G, Selig S, Tarekegn A, Bekele E, Bradman N, Wasser WG, Behar DM, Skorecki K. Missense mutations in the APOL1 gene are highly associated with end stage kidney disease risk previously attributed to the MYH9 gene. *Hum Genet*. 128(3):345–350.
38. Shrestha S, Irvin MR, Taylor KD, Wiener HW, Pajewski NM, Haritunians T, Delaney JA, Schambelan M, Polak JF, Arnett DK, Chen YD, Grunfeld C. A genome-wide association study of carotid atherosclerosis in HIV-infected men. *AIDS*. 24(4):583–592.